# Determining the spatial distribution of soil properties using the environmental covariates and multivariate statistical analysis: a case study in semi-arid regions of Iran

Mojtaba ZERAATPISHEH[1,2,3*], Shamsollah AYOUBI[1], Magboul SULIEMAN[4], Jesús RODRIGO-COMINO[5]

[1] Department of Soil Science, College of Agriculture, Isfahan University of Technology, Isfahan 84156-83111, Iran;
[2] Key Laboratory of Geospatial Technology for the Middle and Lower Yellow River Regions, College of Environment and Planning, Henan University, Kaifeng 475004, China;
[3] Department of Soil Science, Agricultural Sciences and Natural Resources University of Khuzestan, Ahvaz 6341773637, Iran;
[4] Department of Soil and Environment Sciences, Faculty of Agriculture, University of Khartoum, Shambat 13314, Sudan;
[5] Instituto de Geomorfología y Suelos, Department of Geography, University of Málaga, Málaga 29071, Spain

**Abstract:** Natural soil-forming factors such as landforms, parent materials or biota lead to high variability in soil properties. However, there is not enough research quantifying which environmental factor(s) can be the most relevant to predicting soil properties at the catchment scale in semi-arid areas. Thus, this research aims to investigate the ability of multivariate statistical analyses to distinguish which soil properties follow a clear spatial pattern conditioned by specific environmental characteristics in a semi-arid region of Iran. To achieve this goal, we digitized parent materials and landforms by recent orthophotography. Also, we extracted ten topographical attributes and five remote sensing variables from a digital elevation model (DEM) and the Landsat Enhanced Thematic Mapper (ETM), respectively. These factors were contrasted for 334 soil samples (depth of 0–30 cm). Cluster analysis and soil maps reveal that Cluster 1 comprises of limestones, massive limestones and mixed deposits of conglomerates with low soil organic carbon (SOC) and clay contents, and Cluster 2 is composed of soils that originated from quaternary and early quaternary parent materials such as terraces, alluvial fans, lake deposits, and marls or conglomerates that register the highest SOC content and the lowest sand and silt contents. Further, it is confirmed that soils with the highest SOC and clay contents are located in wetlands, lagoons, alluvial fans and piedmonts, while soils with the lowest SOC and clay contents are located in dissected alluvial fans, eroded hills, rock outcrops and steep hills. The results of principal component analysis using the remote sensing data and topographical attributes identify five main components, which explain 73.3% of the total variability of soil properties. Environmental factors such as hillslope morphology and all of the remote sensing variables can largely explain SOC variability, but no significant correlation is found for soil texture and calcium carbonate equivalent contents. Therefore, we conclude that SOC can be considered as the best-predicted soil property in semi-arid regions.

**Keywords:** soil properties; remote sensing data; topographical attributes; multivariate statistical analyses; geographic information systems; land management

*Corresponding author: Mojtaba ZERAATPISHEH (E-mail: zeraatpishem@yahoo.com)

## 1 Introduction

In arid and semi-arid areas, vegetation patterns and pedogenesis are highly affected by extreme climate conditions, and also by the topography (Jia et al., 2017; Cerdà et al., 2018). Flat areas, such as alluvial fans, valleys or lagoons, that are highly exposed to high peaks of insolation, register high average temperatures with prolonged dry seasons that enhance high evapotranspiration rates (Li et al., 2016), the difficulties of organic matter consolidation (Wang et al., 2016), the generation of soil crusts (Singer and Shainberg, 2004) and the reduction of available water for the plants (Amare et al., 2014).

Specifically, in semi-arid landscapes, soil texture, organic carbon content and calcium carbonate content are able to reflect the spatial distributions of soil nutrients and vegetative cover (Shiri et al., 2017; Keshavarzi et al., 2018). Soil particle distribution is considered as a key parameter of soil quality because soil clay content controls soil cation-exchange capacity (Khaledian et al., 2017a; Sulieman et al., 2018), soil aggregate stability (Schjønning et al., 2007), and soil organic carbon preservation (Zeraatpishe and Khormali, 2012). Moreover, soil silt is also highly related to soil erosion and is sensitive to soil particle detachment and transport processes (Ramos et al., 2000). Sand and clay particles determine water content in the field and wilting point capacities, depending on the carbonate and organic carbon contents (Wang et al., 2013). Moreover, soil organic carbon stores and mineral horizons also show a high spatial variation due to the characteristics of topography, slope gradient and curvature (Conforti et al., 2016), and the parent material and type of land use (Sun et al., 2015). Also, based significantly on parent material, landform distributions and hydrological conditions, several scholars have demonstrated that soil calcium carbonate is accumulated, transported and distributed (Khormali et al., 2009; Karchegani et al., 2011; Wilford et al., 2015). In regard to topographic position, Stavi et al. (2008) concluded that calcium carbonate content in soil profiles can be lower on the north-facing hillslope than on the south-facing hillslope in the Negev region of Israel because of the lower evaporation on the north-facing hillslope. However, there is a lack of information about whether the combination of soil texture, organic carbon and carbonate calcium follows common spatial patterns in the same landscape and is conditioned by topographical variables, parent materials and/or vegetation covers.

Cluster analysis (CA) and principal component analysis (PCA) are considered as useful statistical techniques to assess the spatial relationships among soil variables and to obtain a group of factors to explain the variability of the soil system (Shukla et al., 2006). Both CA and PCA are frequently and respectively used as they are capable of detecting similarities among soil samples and/or variables (Tran et al., 2010; Ruggieri et al., 2011). However, they are rarely used in combination previously.

Moreover, in order to check the obtained results from the multivariate statistical analyses with the spatial distribution of the soil properties, several researchers also highlighted the use of geostatistical analysis as the most useful tool (Gribov and Krivoruchko, 2012; Martínez-Murillo et al., 2017). Using remote sensing data and Geographic Information Systems (GISs), techniques such as inverse distance weighting (IDW), ordinary kriging, empirical Bayesian kriging and radial basis functions are widely used in soil sciences and other disciplines such as hydrology, geology and environmental sciences (Sağir and Kurtuluş, 2017; Samsonova et al., 2017).

The objective of this research is to investigate the ability of multivariate statistical analyses (CA and PCA) to distinguish which soil properties follow a clear spatial pattern conditioned by specific environmental characteristics, using the IDW interpolation method, in a semi-arid case study catchment located in central Iran.

## 2 Materials and methods

### 2.1 Study area

The study area is located in the Borujen region, Chaharmahal-Va-Bakhtiari Province, central Iran. This area extends from 31°41′57″ to 32°00′01″N latitude and from 51°02′45″ to 51°19′09″E longitude (Fig. 1), and covers an area of approximately 860 km$^2$ with a mean elevation of about 2200 m a.s.l. (Zeraatpisheh et al., 2017).

The climate conditions can be defined as semi-arid, with mean annual precipitation and annual mean temperature of 255 mm and 10.7°C, respectively. The main parent materials are limestones (Zeraatpisheh et al., 2017; 2019; Ayoubi et al., 2018). Soil moisture and temperature regimes are xeric and mesic, respectively, according to the soil taxonomy (Soil Survey Staff, 2014). The main land uses include irrigated wheat cropping, dryland farming, and pastures. Predominant crops are wheat, barley and alfalfa. In pasture land uses, the main plant species are milkvetch and *Bromus tectorum*. The major landscape units in the study area comprise mountains, hill-lands, piedmonts, and lowlands.
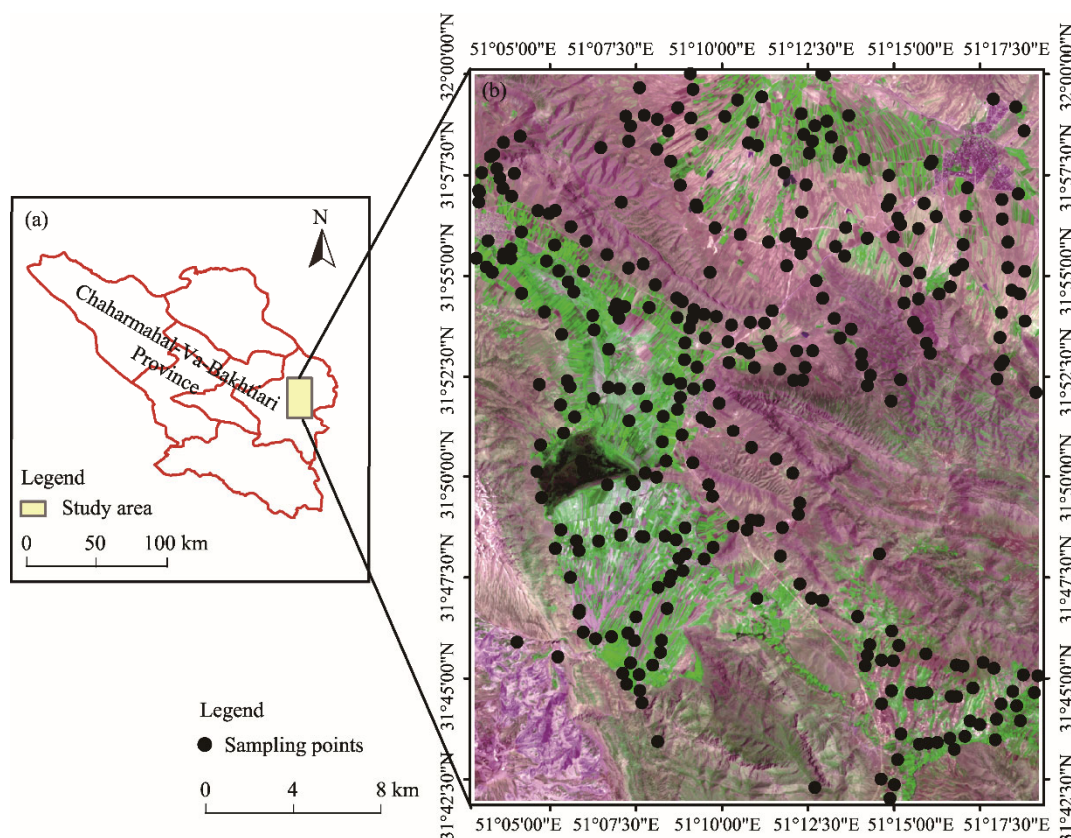
**Fig. 1**   Location of the study area in Chaharmahal-Va-Bakhtiari Province (a) and distributions of soil sampling points (black circles) in the study area based on Landsat ETM (colour composition; b)

## 2.2   Soil sampling method and laboratory analysis

Distributions of 334 soil sampling points acquired by conditioned Latin Hypercube Sampling (cLHS) (Minasny and McBratney, 2006) are shown in Figure 1. Soil sampling to depths of 0–30 cm was performed at each sampling point and Matlab software (Mathworks) was used to estimate the quantitative attributes from the topographical data and remote sensing data (Table 1). Three soil properties were analyzed: soil organic carbon (SOC), calcium carbonate equivalent (CCE) and soil texture (clay, silt, and sand). SOC was determined by a modified Walkley and Black method according to Nelson and Sommers (1996). CCE was measured using the method of Page et al. (1982) and soil particle size distribution was determined by the Bouyoucos hydrometer method (Bouyoucos, 1962).

## 2.3   Soil mapping, remote sensing data and topographical attributes

SOC, CCE and soil texture (clay, silt, and sand) contents were represented in the form of maps using the inverse distance weighting (IDW) interpolation method with ArcMap 10.5 (ESRI, USA). By cross-validation, performance statistics of $R^2$ (coefficient of determination) and RMSE (root mean

square error) were calculated for IDW interpolation to validate which soil property is the best predicted.

Geological units were extracted using the 1:100,000 Borujen Geology Map (1990) and land unit areas by air photointerpretation. Topographical attributes were obtained from a Digital Elevation Model (DEM) with a cell size of 30 m, which was provided by the ASTER GDEM database (ASTER GDEM, 2009). Remote sensing indices were derived from the Landsat Enhanced Thematic Mapper (ETM) acquired in 2008 (U.S. Geological Survey, 2004). Ten topographical attributes including elevation, topographic wetness index (TWI), SAGA (System for Automated Geoscientific Analysis) wetness index (WI), multi-resolution of ridge top flatness index (MrRTF), multi-resolution valley bottom flatness index (MrVBF), curvature (Cur), profile curvature (PrCu), plan curvature (PlCu), aspect (Asp) and slope (Slp) at the same resolution were calculated. Furthermore, five remote sensing auxiliary variables, namely the normalized difference vegetation index (NDVI), ratio vegetation index (RVI), perpendicular vegetation index (PVI), clay index (CI), and soil-adjusted vegetation index (SAVI) were also extracted. The SAGA GIS was employed to derive the environmental factors (Olaya, 2004) (Table 1).

**Table 1** Environmental covariates used in this study

| Environmental factor | Nature of the soil variable derived | Name of factor | Definition | Type |
|---|---|---|---|---|
| Topography | DEM | El | Elevation (m) | Quantitative |
| | | TWI | Topographic wetness index | |
| | | WI | Wetness index | |
| | | MrRTF | Multi-resolution of ridge top flatness index | |
| | | MrVBF | Multi-resolution valley bottom flatness index | |
| | | Cur | Curvature | |
| | | PrCu | Profile curvature | |
| | | PlCu | Plan curvature | |
| | | Asp | Aspect (radians) | |
| | | Slp | Slope (°) | |
| Biogeography | Landsat ETM | NDVI | Normalized difference vegetation index | Quantitative |
| | | RVI | Ratio vegetation index | |
| | | PVI | Perpendicular vegetation index | |
| | | CI | Clay index | |
| | | SAVI | Soil adjusted vegetation index | |
| Landform | Landscape and landform | GEM | Landforms (10 units) | Categorical |
| Geology map | Geologic unit | GEO | Lithological units (12 units) | |

Note: DEM, Digital Elevation Model; ETM, Enhanced Thematic Mapper.

## 2.4 Statistical analysis

Descriptive statistics of soil properties, topographical attributes, and remote sensing data, including minimum, maximum, mean, standard deviation, variance, the coefficient of variation (CV), median, skewness and kurtosis, were determined.

The CA was undertaken for the different geological and landform units, and average linkages between groups for dissimilarities in the variables were calculated. Then, hierarchical clustering series using Euclidean distance and squared Euclidean distance methods for geology and landform units were extracted from the dataset, respectively. These techniques have been applied to assess diverse natural source variations in soil properties and groups of soil types among similar clusters (Massart and Kaufman, 1983).

Furthermore, the PCA was carried out to identify distribution and relationships of soil texture, SOC, and CCE with environmental factors. The PCA allowed grouping of similar variables into

dimensions, without distinguishing independent and dependent variables. Firstly, the Pearson's correlation was conducted to check the linear correlations among variables. Secondly, the variables were centred and normalized. Then, a Kaiser-Mayer-Olkin (KMO) test of sphericity was conducted and the coefficients and determinants were observed to verify the assumptions. The orthogonal rotation method (Varimax) and the correlation matrix using factors with eigenvalues >1 were performed. The variables used to achieve our main goal should model the phenomenon observed (soil properties, topographical attributes and remote sensing data) in the closest possible way, identifying specifically the most significant variables that condition soil distribution and formation (Malinowski, 2002; Shukla et al., 2006). Factor loading indicates the correlation between variable and underlying common factor. The highly loading variables can be used to propose a possible common underlying factor that links variables together within each factor. The signs of the factor loadings provide information on how these variables are related and when they are representing the common factors.

Our dataset included twenty environmental factors, which were grouped as follows: (i) topographical attributes, namely El (elevation), Asp, TWI, WI, MrVBF, MrRTF, Slp, Cur, PlCu, and PrCu; (ii) remote sensing variables, namely NDVI, RVI, PVI, CI, and SAVI; and (iii) soil properties of soil particle distribution (clay, silt and sand), SOC content and CCE content. All statistical analyses were performed using SPSS 23.0 (IBM Corp, 2015, USA).

## 3 Results

### 3.1 Soil properties, topographical attributes and remote sensing data

In Figures 2–4, soil properties, geology, and landforms in the study area are represented in the form of maps. Moreover, descriptive statistics of soil properties, topographical attributes, and remote sensing data are summarized in Table 2. Also, total areas of each geological and landform units are summarized in Tables 3 and 4, respectively.
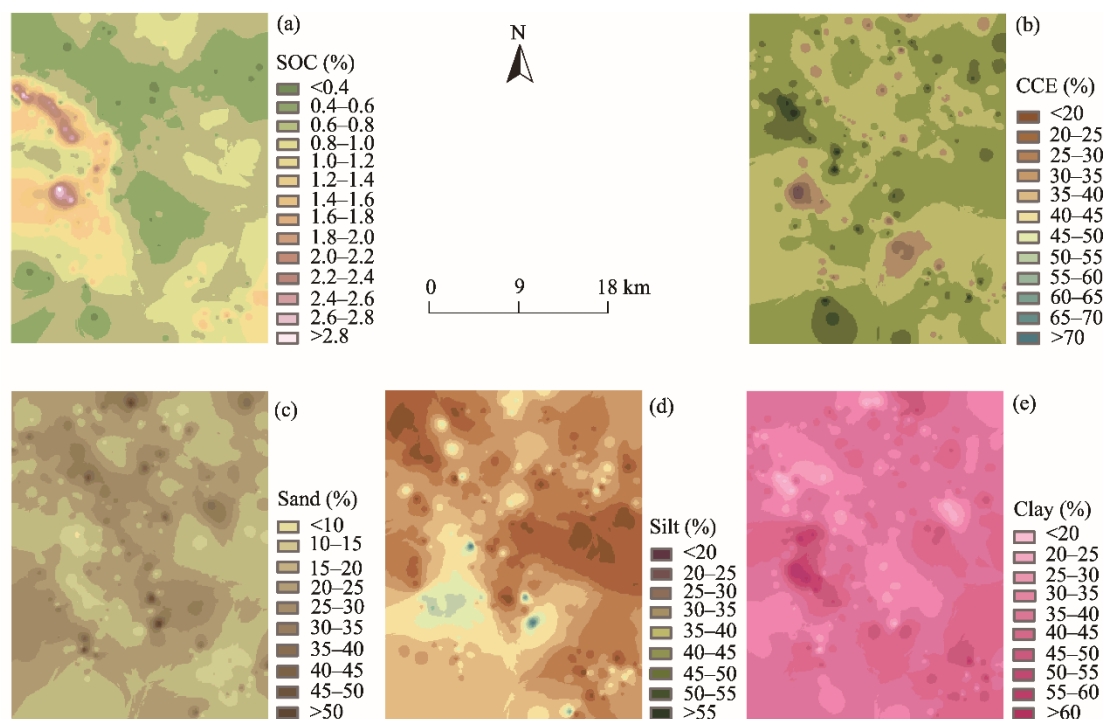


**Fig. 2** Distributions of soil properties in the study area using inverse distance weighting (IDW) interpolation method. (a), soil organic carbon (SOC); (b), calcium carbonate equivalent (CCE); (c), sand content; (d), silt content; (e), clay content.
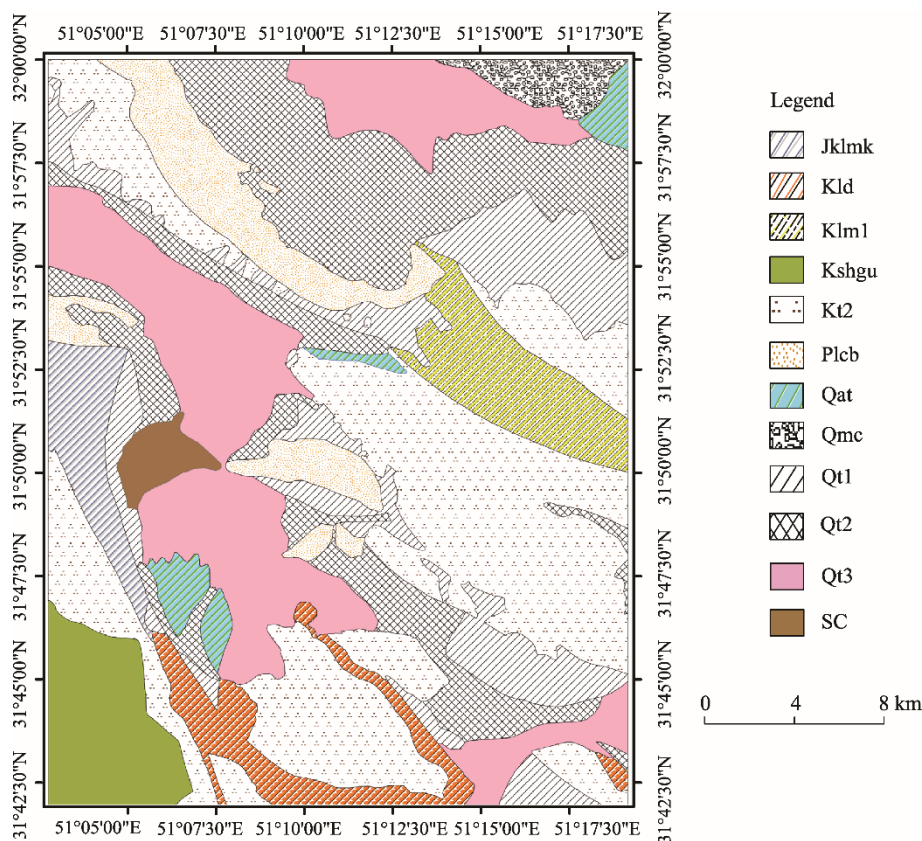
**Fig. 3** Geological units in the study area. Jklmk, mixed deposit of limestone, marl and shale; Kld, massive dark grey limestone, marl and shale; Klm1, marl and limestone; Kshgu, marl and rare sandstone; Kt2, dark grey massive limestone; Plcb, thick-bedded conglomerate with marl; Qat, river channel and recent alluvium; Qmc, grey marl, conglomerate and lake deposit; Qt1, older terrace and alluvial fan; Qt2, young terrace and alluvial fan; Qt3, silty clay flat; SC, silty clay flat (Zeraatpisheh et al., 2017). The abbreviations are the same as in Figure 5.

Dark grey massive limestone and young terrace and alluvial fan units occupy 26.4% and 18.8% of the total area, respectively (Table 3). The recent and youngest parent materials are distributed in the low altitudes, mostly from the centre to the northeast of the study area. Moreover, the old parent materials are distributed in the high altitudes. As shown in Figure 4 and Table 4, a total of ten landform units are obtained, and the areas of Mo1 (rock outcrop) and Pi1 (alluvial fan) are the largest, with the total area percentages of 33.9% and 30.4%, respectively. Mo1, Hi1 (eroded hill land), and Hi2 (developed hill land) landforms cover the southwest, the east and some part of the centre in the study area. Pi1 and Pi5 (piedmont plain) landforms mostly are located in the centre with the youngest parent materials.

The highest SOC contents are observed from the centre to the southwest and the west of the study area (Fig. 2a), coinciding with Pl1 (wetland), Pl2 (lagoon), Pi1 and Pi4 (river plain) landform units (Fig. 4). However, the lowest SOC values are found in the northern, north-western and north-eastern parts, particularly along the Mo1, Hi1 and Hi2 landform units (Fig. 4).

From Figure 2b we can find that the lowest CCE values are located in the northern parts and there are only three small patches with values higher than 60.0%. Soils with the highest clay contents are in the lowest altitudes, which correspond to the Pl1 and Pl2 landform units (Figs. 2e and 4).

Regarding descriptive statistics (Table 2), SOC content ranges from 0.3% to 2.9% (CV=58.1%). Related to soil texture, clay content ranges from 18.0% to 6.01% (CV=20.1%), silt from 20.0% to 60.0% (CV=15.8%) and sand from 8.0% to 54.0% (CV=35.5%). Furthermore, CCE content shows values from 11.0% to 73.0% (CV=31.4%). The skewness values of CCE, clay and silt contents are

less than 0.54, and the kurtosis values are less than 0.88. It should be noted that SOC content shows the highest values of skewness and kurtosis among all soil properties.
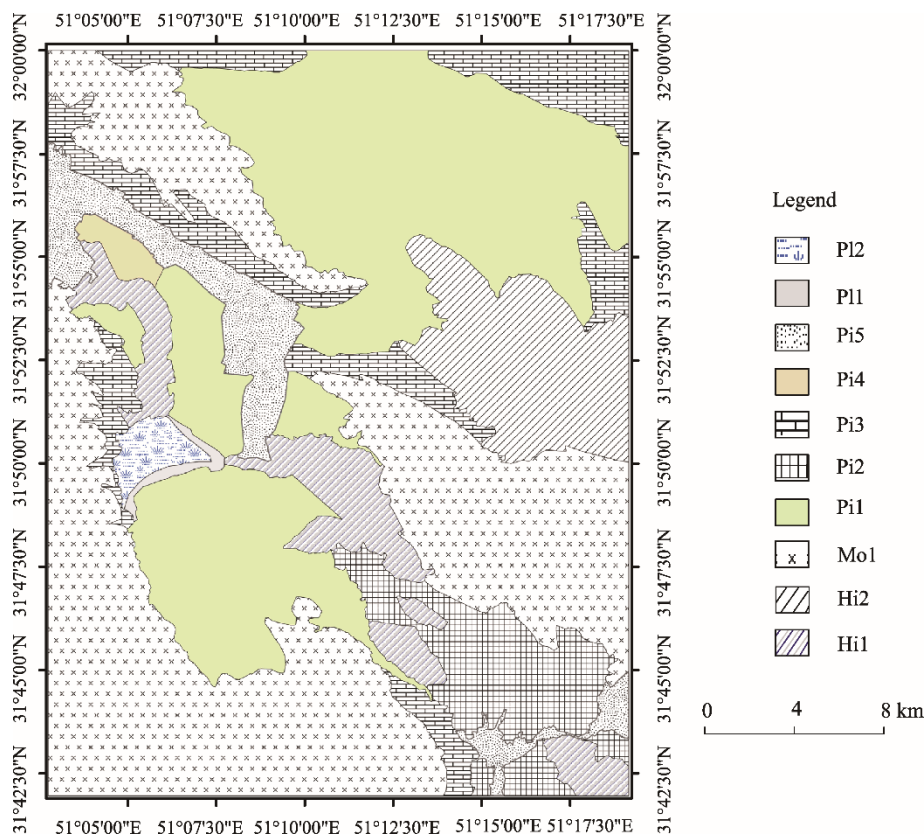
**Fig. 4** Landform units in the study area. Pl2, lake (lagoon); Pl1, wetland; Pi5, piedmont plain; Pi4, river plain; Pi3, pediment; Pi2, dissected alluvial fan; Pi1, alluvial fan; Mo1, rock outcrop; Hi2, developed hill land; Hi1, eroded hill land. The abbreviations are the same as in Figure 4.

Table 5 provides the results of $R^2$ and RMSE indices to demonstrate which soil property registers the best validation using IDW. The results show that SOC registers the highest $R^2$ value (0.62) and the lowest RMSE value (0.30), while silt records the lowest $R^2$ value (0.10).

### 3.2 Influences of parent materials and landform units on soil properties

Figure 5 shows the dendrogram of geology units, where three highly differentiated clusters showing high similarities can be noted. From the top to the bottom of the dendrogram, the first group (Cluster 1) contains the oldest parent materials (Cretaceous) composed of Kt2 (dark grey massive limestone), Plcb (thick-bedded conglomerate with marl), Jklmk (mixed deposit of limestone, marl and shale), Kld (massive dark grey limestone, marl and shale), and Klm1 (marl and limestone). Cluster 2 is grouped with the quaternary parent materials of Qmc (grey marl, conglomerate and lake deposit), Qt2 (young terrace and alluvial fan), Qat (river channel and recent alluvium), Qt1 (older terrace and alluvial fan), Qt3 (silty clay flat), and SC (silty clay flat). Moreover, it consists of two sub-clusters that are separated by other two different geological formations of Qt3 and SC. Cluster 3 contains one geological unit characterized by Kshgu. We observe that Cluster 1 contains the majority of lithologies with a lower amount of SOC and clay contents compared with the surrounding areas. Cluster 2 contains the lithologies that comprise soils with the highest SOC content and the lowest sand and silt contents.

A hierarchical CA is employed to detect dissimilarities among landform units. A dendrogram of landform units is generated to group the sampling points into two statistically significant clusters (Fig. 6). The first cluster contains soils that are situated on piedmonts, eroded hillslopes and lowland landscapes that comprise moderately to highly developed soils. The second cluster shows soils

**Table 2** Descriptive statistics of soil properties, topographical attributes and remote sensing variables

| Factor | Min. | Max. | Mean | SD | Variance | CV | Median | Ske. | Kurtosis |
|---|---|---|---|---|---|---|---|---|---|
| SOC (%) | 0.30 | 2.90 | 0.90 | 0.50 | 0.30 | 58.10 | 0.70 | 2.13 | 4.94 |
| CCE (%) | 11.00 | 73.00 | 32.50 | 10.20 | 104.00 | 31.40 | 32.00 | 0.54 | 0.63 |
| Clay (%) | 18.00 | 61.00 | 37.90 | 7.60 | 58.00 | 20.10 | 38.00 | 0.11 | 0.47 |
| Silt (%) | 20.00 | 60.00 | 40.20 | 6.30 | 40.20 | 15.80 | 40.00 | −0.01 | 0.88 |
| Sand (%) | 8.00 | 54.00 | 21.90 | 7.80 | 60.30 | 35.50 | 21.00 | 1.11 | 1.75 |
| El (m) | 2147.90 | 2710.90 | 2268.00 | 74.90 | 5616.00 | 3.30 | 2250.00 | 1.89 | 5.63 |
| TWI | 13.71 | 19.13 | 14.79 | 1.18 | 1.40 | 8.00 | 14.30 | 1.36 | 1.36 |
| WI | 9.38 | 15.07 | 12.07 | 1.28 | 1.64 | 10.60 | 12.06 | −0.05 | −1.11 |
| MrVBF | 0.00 | 5.95 | 1.81 | 1.54 | 2.38 | 85.46 | 1.39 | 0.77 | −0.46 |
| MrRTF | 0.00 | 3.98 | 0.76 | 0.90 | 0.82 | 119.28 | 0.34 | 1.45 | 1.22 |
| Cur | −0.010 | 0.010 | 0.000 | 0.001 | 0.000 | −488.350 | 0.000 | −1.200 | 8.090 |
| Asp (radians) | 0.54 | 5.88 | 3.19 | 1.38 | 1.90 | 43.13 | 3.29 | −0.05 | −1.00 |
| PlCu | −0.002 | 0.002 | 0.000 | 0.001 | 0.000 | 869.400 | 0.000 | −0.170 | 1.610 |
| PrCu | −0.008 | 0.005 | 0.000 | 0.001 | 0.000 | −321.610 | 0.000 | −1.400 | 12.190 |
| Slp (°) | 0.00 | 0.41 | 0.07 | 0.06 | 0.00 | 89.91 | 0.05 | 2.26 | 8.42 |
| NDVI | −0.11 | 0.47 | 0.04 | 0.13 | 0.02 | 351.54 | −0.01 | 1.22 | 0.80 |
| RVI | −0.004 | 0.001 | 0.000 | 0.001 | 0.000 | −401.220 | 0.000 | −1.070 | 0.320 |
| PVI | 6.61 | 135.60 | 73.49 | 30.07 | 904.45 | 40.92 | 78.95 | −0.51 | −0.64 |
| CI | 1.11 | 1.85 | 1.24 | 0.13 | 0.02 | 10.12 | 1.19 | 1.85 | 3.90 |
| SAVI | −0.17 | 0.70 | 0.05 | 0.19 | 0.04 | 351.54 | −0.01 | 1.22 | 0.80 |

Note: Min., minimum; Max., maximum; SD, standard deviation; CV, coefficient of variation; Ske., Skewness; SOC, soil organic carbon; CEE, calcium carbonate equivalent. The unit measurements of minimum, maximum, mean, SD, variance, CV and median are the same with the corresponding environmental factor, while skewness and kurtosis are dimensionless parameters.

**Table 3** Characteristics of geological units in the study area

| Geology | Age | Geology formation | Area (km$^2$) | Percentage of total area (%) |
|---|---|---|---|---|
| SC | Young quaternary | Silty clay flat | 9.6 | 1.1 |
| Qat | Young quaternary | River channel and recent alluvium | 15.9 | 1.8 |
| Qt3 | Quaternary | Silty clay flat | 151.0 | 17.5 |
| Qt2 | Quaternary | Young terrace and alluvial fan | 162.1 | 18.8 |
| Qt1 | Quaternary | Older terrace and alluvial fan | 94.7 | 11.0 |
| Qmc | Early quaternary | Grey marl, conglomerate and lake deposit | 12.3 | 1.4 |
| Plcb | Pliocene | Thick-bedded conglomerate with marl | 57.4 | 6.7 |
| Kt2 | Late Cretaceous | Dark grey massive limestone | 228.1 | 26.4 |
| Klm1 | Late Cretaceous | Marl and limestone | 38.4 | 4.4 |
| Kshgu | Late Cretaceous | Marl and rare sandstone | 40.1 | 4.6 |
| Jklmk | Early Cretaceous | Mixed deposit of limestone, marl and shale | 24.3 | 2.8 |
| Kld | Early Cretaceous | Massive dark grey limestone, marl and shale | 30.1 | 3.5 |

**Table 4** Characteristics of landform units in the study area

| Landscape | Landform | Landform code | Area (km$^2$) | Percentage of total area (%) | Definition |
|---|---|---|---|---|---|
| Hill | Eroded hill land | Hi1 | 39.2 | 4.5 | Continuous hills with a high altitude |
| | Developed hill land | Hi2 | 62.9 | 7.3 | Active fan, cultivated plain in the low altitudes |
| Mountain | Rock outcrop | Mo1 | 292.9 | 33.9 | Eroded rock surface, shallow soils |
| Piedmont | Alluvial fan | Pi1 | 262.3 | 30.4 | Dissected and undulating red alluvial fan |
| | Dissected alluvial fan | Pi2 | 52.6 | 6.1 | Shallow soil, colluvial material |
| | Pediment | Pi3 | 79.6 | 9.2 | Low drainage, young terraces |
| | River plain | Pi4 | 5.7 | 0.7 | Young terraces, moderate depth soils |
| | Piedmont plain | Pi5 | 57.4 | 6.7 | Low drainage, wetness |
| Lowland | Wetland | Pl1 | 3.0 | 0.4 | Covered by hydrophilic plants |
| | Lake (Lagoon) | Pl2 | 8.2 | 1.0 | Single and low topography hills |

**Table 5**  Validation criteria for inverse distance weighting (IDW) interpolation method among soil properties

| Soil property | $R^2$ | RMSE |
|---|---|---|
| SOC | 0.62 | 0.30 |
| CCE | 0.35 | 8.25 |
| Clay | 0.35 | 6.17 |
| Silt | 0.10 | 6.27 |
| Sand | 0.16 | 7.21 |

Note: $R^2$, coefficient of determination; RMSE, root mean square error.

originated from Mo1 and Hi2 landform units that are mostly characterized by rock outcrops and eroded rock surfaces with shallow to very shallow soil depths. The CA allows us to distinguish landforms that register soils with a higher amount of SOC and clay contents (Cluster 1); meanwhile, Cluster 2 contains soils with lower SOC content values and clay contents.
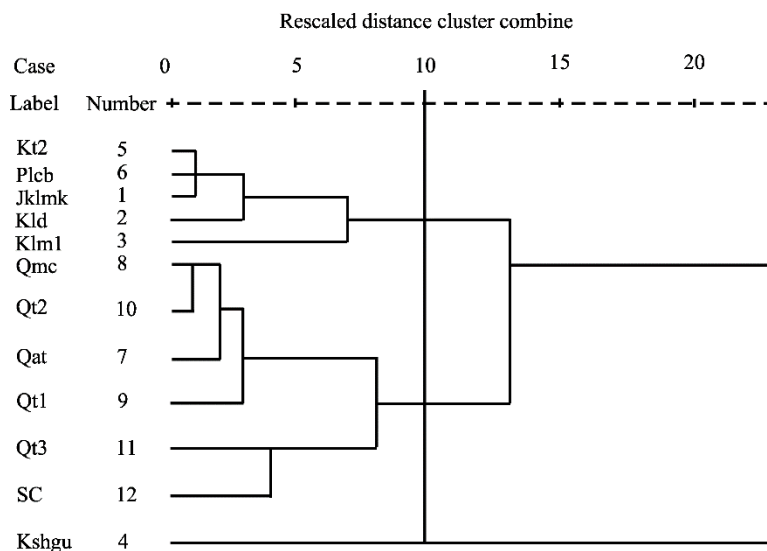


**Fig. 5**  Cluster analysis of geology units in the study area
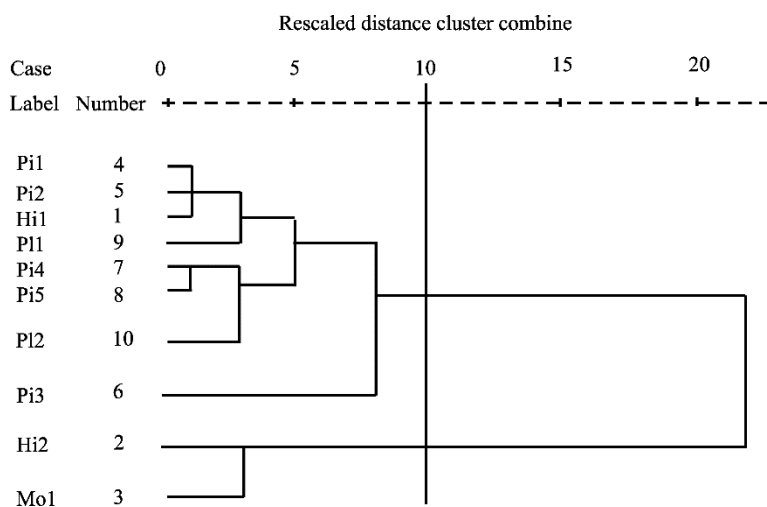


**Fig. 6**  Cluster analysis of landform units in the study area

## 3.3    Correlations among soil properties, topographical attributes and remote sensing variables

Pearson's correlation analysis between soil properties and environmental factors are presented in Table 6. Absolute correlation values higher than 0.40 are considered significant. In this context, a

significant correlation exits between SOC and MrVBF and between SOC and all the remote sensing variables. As a result, SOC can be considered as a potential soil indicator to confirm the vegetation development and abundance obtained from the remote sensing analysis. Clay and silt contents show significant negative correlations with sand content, with correlation coefficients of –0.63 and –0.43, respectively. Moreover, CCE does not show significant correlations with any soil property or environmental factor (topographical and remote sensing variables). The El shows slightly positive correlations with Slp ($r$=0.45, $P$<0.05) and TWI ($r$=0.41, $P$<0.05). TWI also shows negative correlations with MrRTF and MrVBF. Additionally, WI shows a positive correlation with MrVBF ($r$=0.52, $P$<0.05) and a negative one with Slp ($r$= –0.50, $P$<0.05). As expected, curvature also exhibits positive correlations with PlCu ($r$=0.68, $P$<0.05) and PrCu ($r$=0.94, $P$<0.05). Finally, it is interesting to observe that soil properties show the highest correlations with all the remote sensing variables, but surprisingly, NDVI does not show any significant correlation with some of them (Table 6).

### 3.4　Principal component analysis (PCA)

The results of PCA show five main components to explain the variability of the selected soil properties (Table 7). It shows that the first two components (PCA 1 and PCA 2) can explain more than 42.0% of the total variability. Then, when including the other three components, a total of 73.3% is explained. In Table 8, the obtained components are added and the factors are identified. The diagram of factors obtained to explain the geomorphological influence on soil properties is shown in Figure 7. The commonalities of twenty factors considered in this study indicate that these five components explain a large part of the variation in most of the measured factors (Table 8).

Component 1 is able to explain about 27.3% of the total variance and it is highly contributed to by SOC and all of the remote sensing variables (Tables 7 and 8). Once again, it is remarkable to observe the strong relationships between SOC and remote sensing variables. Component 2 can explain about 15.1% of the total variance, with significant contributions from the landform covariates related to the hillslope morphology and wetness indices such as El, TWI, WI, MrVBF, and Slp. Component 3 explains about 14.5% of the total variance, being highly contributed to by the curvature and wetness indices such as TWI, WI, MrRTF, Cur, PlCu, and PrCu. Component 4 explains about 8.4% of the total variance and is dominantly contributed to by clay and sand contents. Lastly, component 5 accounts for about 7.9% of the total variance, with significant contributions from CCE, and silt and sand contents. In general, we conclude that the first component is highly contributed to by remote sensing variables and SOC, and the second and third components by wetness indices and the morphology of the hillslope. Conversely, the fourth and fifth components primarily include soil properties and do not relate strongly with topographical attributes and remote sensing data (Fig. 7).

## 4　Discussion

Among all soil properties, SOC shows the highest CV value of 58.1%, which indicates a broad range of SOC values across the study area. Wilding (1985) categorized CV values into three classes of high variability (CV≥35%), moderate variability (15%≤CV<35%) and low variability (CV<15%). Also, the high skewness and kurtosis values of SOC confirm the vast distribution of SOC in the study area (Rosemary et al., 2017). Therefore, this implies the importance of knowing which kind of spatial pattern can SOC follow under different environmental conditions, and this needs to be further clarified and determined. The high variation of SOC could also be due to the land use practices. In semi-arid areas, agricultural activities can lead to a significant increase in SOC in comparison to bare lands. In China, for example, Liao et al. (2015) concluded that agricultural intensification increases crop productivity and crop residues, and also highly affects SOC values and carbon sequestration.

The CVs for the other soil properties are more moderate, coinciding with the results in other semi-arid areas with non-active land use changes. Also, the predominance of parent materials can be another driving factor as other researchers confirmed (Anderson, 1988; Bruand and Tessier, 2000; Orgill et al., 2017). Therefore, the moderate variability of CCE content might be explained because all soils are formed on limestone and calcareous materials. It is important to highlight that high levels of CCE in the southwest of the study area could be affected by soil crusts, abandonment of

**Table 6** Pearson's correlation analysis among soil properties, topographical attributes and remote sensing variables

| | SOC | CCE | Clay | Silt | Sand | El | TWI | WI | MrVBF | MrRTF | Cur | Asp | PlCu | PrCu | Slp | NDVI | RVI | PVI | CI | SAVI |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SOC | 1.00 | 0.24 | 0.02 | 0.13 | -0.12 | -0.31 | -0.27 | 0.22 | 0.51* | 0.07 | 0.03 | -0.10 | -0.02 | 0.04 | -0.29 | 0.63* | -0.65* | -0.61* | 0.67* | 0.63* |
| CCE | | 1.00 | -0.01 | -0.25 | 0.22 | -0.23 | -0.26 | 0.08 | 0.31 | 0.18 | 0.06 | -0.05 | 0.06 | 0.04 | -0.14 | 0.25 | -0.25 | -0.17 | 0.29 | 0.25 |
| Clay | | | 1.00 | -0.39 | -0.66* | -0.12 | -0.17 | -0.08 | 0.13 | 0.18 | 0.12 | 0.00 | 0.04 | 0.13 | -0.11 | 0.08 | -0.11 | -0.10 | 0.09 | 0.08 |
| Silt | | | | 1.00 | -0.43* | 0.02 | 0.06 | 0.05 | -0.03 | -0.13 | -0.06 | 0.03 | -0.06 | -0.05 | 0.04 | 0.14 | -0.12 | -0.15 | 0.12 | 0.14 |
| Sand | | | | | 1.00 | 0.10 | 0.11 | 0.04 | -0.11 | -0.07 | -0.06 | -0.03 | 0.01 | -0.08 | 0.08 | -0.19 | 0.20 | 0.22 | -0.18 | -0.19 |
| El | | | | | | 1.00 | 0.41* | -0.13 | -0.39 | -0.27 | -0.11 | 0.03 | -0.02 | -0.13 | 0.45* | -0.32 | 0.32 | 0.26 | -0.29 | -0.32 |
| TWI | | | | | | | 1.00 | -0.05 | -0.53* | -0.41* | -0.34 | 0.10 | -0.32 | -0.28 | 0.59* | -0.22 | 0.22 | 0.12 | -0.21 | -0.22 |
| WI | | | | | | | | 1.00 | 0.52* | -0.03 | -0.35 | -0.04 | -0.37 | -0.27 | -0.50* | 0.13 | -0.13 | -0.13 | 0.14 | 0.13 |
| MrVBF | | | | | | | | | 1.00 | 0.19 | -0.03 | -0.03 | -0.17 | 0.04 | -0.64 | 0.38 | -0.38 | -0.34 | 0.39 | 0.38 |
| MrRTF | | | | | | | | | | 1.00 | 0.33 | -0.05 | 0.30 | 0.28 | -0.32 | 0.14 | -0.14 | -0.07 | 0.13 | 0.14 |
| Cur | | | | | | | | | | | 1.00 | -0.05 | 0.68* | 0.94* | -0.19 | 0.05 | -0.05 | 0.00 | 0.03 | 0.05 |
| Asp | | | | | | | | | | | | 1.00 | -0.08 | -0.02 | 0.06 | -0.06 | 0.07 | 0.06 | -0.04 | -0.06 |
| PlCu | | | | | | | | | | | | | 1.00 | 0.39 | 0.02 | 0.01 | -0.01 | 0.06 | 0.01 | 0.01 |
| PrCu | | | | | | | | | | | | | | 1.00 | -0.25 | 0.06 | -0.05 | -0.02 | 0.04 | 0.06 |
| Slp | | | | | | | | | | | | | | | 1.00 | -0.21 | 0.21 | 0.15 | -0.20 | -0.21 |
| NDVI | | | | | | | | | | | | | | | | 1.00 | -0.99* | -0.95* | 0.96* | 1.00* |
| RVI | | | | | | | | | | | | | | | | | 1.00 | 0.96* | -0.95* | -0.99* |
| PVI | | | | | | | | | | | | | | | | | | 1.00 | -0.88* | -0.95* |
| CI | | | | | | | | | | | | | | | | | | | 1.00 | 0.96* |
| SAVI | | | | | | | | | | | | | | | | | | | | 1.00 |

Note: *, correlation significance at $P<0.05$ level.

**Table 7**  Explained variance of the first five components of principal component analysis (PCA)

| Component | Eigenvalue | Percentage of variance (%) | Accumulated percentage of variance (%) |
|---|---|---|---|
| 1 | 5.466 | 27.3 | 27.3 |
| 2 | 3.019 | 15.1 | 42.4 |
| 3 | 2.908 | 14.5 | 57.0 |
| 4 | 1.686 | 8.4 | 65.4 |
| 5 | 1.575 | 7.9 | 73.3 |

**Table 8**  Eigenvectors of the selected variables for the first five components of PCA

| Factor | Component | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| SOC | 0.677# | 0.342 | −0.021 | 0.077 | −0.031 |
| CCE | 0.257 | 0.231 | 0.051 | 0.217 | 0.597# |
| Clay | 0.041 | 0.095 | 0.077 | −0.943# | 0.263 |
| Silt | 0.177 | 0.001 | −0.026 | 0.143 | −0.902# |
| Sand | −0.185 | −0.094 | −0.055 | 0.807# | 0.478# |
| El | −0.252 | −0.549# | −0.117 | 0.076 | −0.103 |
| TWI | −0.109 | −0.683# | −0.401# | 0.064 | −0.119 |
| WI | 0.060 | 0.605# | −0.526# | 0.137 | −0.081 |
| MrVBF | 0.296 | 0.807# | −0.142 | −0.025 | 0.082 |
| MrRTF | 0.052 | 0.370 | 0.432# | −0.141 | 0.225 |
| Cu | 0.004 | 0.090 | 0.952# | −0.009 | −0.020 |
| As | −0.064 | −0.080 | −0.084 | −0.097 | −0.080 |
| PlCu | −0.001 | −0.084 | 0.785# | 0.055 | 0.065 |
| PrCu | 0.005 | 0.152 | 0.833# | −0.037 | −0.055 |
| Sl | −0.069 | −0.888# | −0.095 | 0.012 | 0.015 |
| NDVI | 0.979# | 0.132 | 0.036 | −0.028 | 0.014 |
| RVI | −0.979# | −0.136 | −0.033 | 0.047 | −0.033 |
| PVI | −0.954# | −0.062 | 0.027 | 0.066 | 0.025 |
| Cl | 0.956# | 0.135 | 0.017 | −0.024 | 0.043 |
| SAVI | 0.979# | 0.132 | 0.036 | −0.028 | 0.014 |

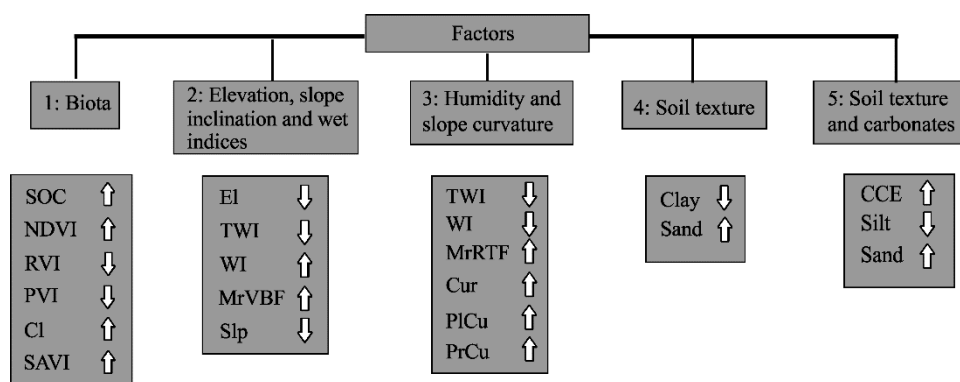Note: # means that values were obtained for each factor.



**Fig. 7**  Diagram of factors obtained to explain the geomorphological influence on soil properties. SOC, soil organic carbon; NDVI, normalized difference vegetation index; RVI, ratio vegetation index; PVI, perpendicular vegetation index; CI, clay index; SAVI, soil-adjusted vegetation index; El, elevation; TWI, topographic wetness index; WI, wetness index; MrVBF, multi-resolution valley bottom flatness index; MrRTF, multi-resolution of ridge top flatness index; Slp, slope; Cur, curvature; PlCu, plan curvature; PrCu, profile curvature; CEE, calcium carbonate equivalent. The upper arrows indicate the increasing trend while the down arrows indicate the decreasing trend.

agricultural activities, and water infiltration (Morin and Van Winkel, 1996; Castillo-Monroy et al., 2010). The locations with high contents of CCE frequently coincide with areas being of low organic matter content and exhibit weak correlations with remote sensing data. It is well-known that carbonates can block organic matter development by blocking the nitrogen interchange (Wang et al., 2016).

However, there is not a common consensus whether parent material or landform could be the main driving factor to explain pedogenesis and vegetation development (Tajik et al., 2012; Mehnatkesh et al., 2013). Our research question also agrees with other studies conducted in semi-arid areas. For example, based on rainfall simulations in semi-arid regions in Spain, Martínez-Hernández et al. (2017) concluded that different parent materials define different hydrological processes at the pedon scale. Differences in parent materials such as marls, sandstones, colluvium and slates can impact on pedogenesis, but those above mentioned researches also remarked that Slp and antecedent soil moisture also play an important role. The results of CA for landform and geology units in the present study confirm that soil properties could be classified into similar groups (clusters). However, by using the qualitative assessment, we observe that soil properties do also coincide with the IDW maps. Thus, it is confirmed that soil mapping techniques and statistical analysis have to be combined (Rodrigo-Comino et al., 2018), but it should also be aware of the parent materials which, as found in our study, control the spatial variation of CCE contents.

With the use of IDW maps, we confirm that this interpolation method is the best tool to get an overview of the studied area (Li and Heap, 2011). In this study, SOC registers the highest $R^2$ and the lowest RMSE values, confirming that this factor obtains similar conclusion as obtained from CA and PCA (high correlation with the spatial variables). However, we acknowledge that there is no assessment of prediction errors from which to produce what are known as "bulls eyes" or higher values near the observed location (http://desktop.arcgis.com/en/arcmap/10.3/tools/geostatistical-analyst-toolbox/idw.htm). To solve this problem, researchers such as Kravchenko and Bullock (1999) added the number of the closest neighbouring samples using cross-validation to compare the results obtained with a different number of the closest neighbouring samples. In future research studies, it should be interesting to test different interpolation methods such as ordinary kriging, empirical Bayesian kriging or radial basis functions to observe if other soil properties do or do not coincide with SOC as the best soil property predictor using multivariate statistical analyses.

Several previous studies such as Dwivedi and Sreenivas (1998) and Metternicht and Zinck (2003) demonstrated that PCA based on the correlation matrix is an effective approach in discriminating soils in arid and semi-arid regions. In our study, results of PCA show that the first five components are able to describe more than 73.3% of the total variability in soil properties. SOC is highly correlated with remote sensing variables. We then conclude the same finding with above mentioned previous studies: an increase of vegetation cover would demonstrate a drastic improvement of SOC. The center to the west of the study area also contains clay to clay loam soils, which confirms that soils can be more fertile and capable of being colonized by natural plants or crops. Regarding the obtained clusters, high SOC contents coincide with Qt3 and SC lithology, and Pl2, Pi1 and Pi4 landforms. These findings demonstrate that parent materials and landform characteristics can show well-conserved SOC levels and vegetative covers as well (Taghizadeh-Mehrjardi et al., 2016). Consequently, lithology and landforms might be considered as indicators of soil fertility. Our findings are agreement with those of Khaledian et al. (2017b), who reported that PCA can be widely used to discriminate soil indicators taking into account spatial variables.

However, in our study, among all the topographic attributes, only MrVBF is significantly correlated with SOC content. The low correlation of topographic attributes with SOC content could be ascribed to the low-relief differences (about 600 m), the extreme semi-arid climate conditions and the high amounts of CCE content. Furthermore, there are slight correlations among silt, sand, El, TWI, WI, MrVBF and all of the remote sensing variables. Alijani and Sarmadian (2014) also confirmed that, in the semi-arid and arid regions in Iran, CCE is positively correlated with some topographical indices such as WI and Cur. Therefore, it could be possible that these parameters can be used to model CCE at larger scales.

ChinaXiv

chinaXiv:201909.00010v1

Based on the findings of this study, we demonstrate that the combination of thematic maps of geology and landforms, topographical attributes and remote sensing data with CA and PCA approaches are able to distinguish clear spatial patterns only for specific soil properties on specific parent materials and landforms. Therefore, further researches need to be done to find the key parameters and factors in other specific areas, which are essential to designing correct and relevant solutions and appropriate land management plans for sustainable development in semi-arid regions.

## 5    Conclusions

This study aims to investigate the ability of multivariate statistical analyses, in combination with topographical attributes and remote sensing data, to identify the spatial distribution of soil properties. Cluster analysis reveals that higher SOC and clay contents are related to flat areas characterized by young quaternary deposits and alluvial fans. Conversely, massive limestones and conglomerate deposits in eroded and consolidated hills and rock outcrops with steep slopes register the lower content of SOC and fine particles, and the higher CCE content. Results of PCA show that the first five components are able to describe more than 73.3% of the total variability in soil properties. Environmental factors such as hillslope morphology and all remote sensing variables can largely explain the variability of SOC, but no significant correlation is found for soil texture and CCE content. Therefore, we conclude that SOC can be the best-predicted soil property because it follows a specific spatial pattern in semi-arid regions conditioned by parent material, hillslope morphology and vegetative cover.

## Acknowledgements

## References

Alijani Z, Sarmadian F. 2014. The role of topography in changing of soil carbonate content. Indian Journal Science Research, 6(1): 263–271.

Amare T, Zegeye A D, Yitaferu B, et al. 2014. Combined effect of soil bund with biological soil and water conservation measures in the northwestern Ethiopian highlands. Ecohydrology & Hydrobiology, 14(3): 192–199.

Anderson D W. 1988. The effect of parent material and soil development on nutrient cycling in temperate ecosystems. Biogeochemistry, 5(1): 71–97.

Ayoubi S, Mokhtari J, Mosaddeghi M R, et al. 2018. Erodibility of calcareous soils as influenced by land use and intrinsic soil properties in a semiarid region of central Iran. Environmental Monitoring and Assessment, 190(4): 192.

Bouyoucos G J. 1962. Hydrometer method improved for making particle size analyses of soils. Agronomy Journal, 54(5): 464–465.

Bruand A, Tessier D. 2000. Water retention properties of the clay in soils developed on clayey sediments: significance of parent material and soil history. European Journal of Soil Science, 51(4): 679–688.

Castillo-Monroy A P, Maestre F T, Delgado-Baquerizo M, et al. 2010. Biological soil crusts modulate nitrogen availability in semi-arid ecosystems: insights from a Mediterranean grassland. Plant and Soil, 333(1–2): 21–34.

Cerdà A, Rodrigo-Comino J, Novara A, et al. 2018. Long-term impact of rainfed agricultural land abandonment on soil erosion in the Western Mediterranean basin. Progress in Physical Geography: Earth and Environment, 42(2): 202–219.

Conforti M, Lucà F, Scarciglia F, et al. 2016. Soil carbon stock in relation to soil properties and landscape position in a forest ecosystem of southern Italy (Calabria region). Catena, 144: 23–33.

Dwivedi RS, Sreenivas K. 1998. Delineation of salt-affected soils and waterlogged areas in the Indo-Gangetic plains using IRS-1C LISS-III data. International Journal of Remote Sensing, 14: 2739–2751.

Gribov A, Krivoruchko K. 2012. New flexible non-parametric data transformation for Trans-Gaussian Kriging. In: Abrahamsen P, Hauge R, Kolbjørnsen O. Geostatistics Oslo. Dordrecht: Springer, 51–65.

IBM Corp. Released 2015. IBM SPSS Statistics for Windows, Version 23.0. Armonk: IBM Corp.

Jia F Q, Tiyip T, Wu N, et al. 2017. Characteristics of soil seed banks at different geomorphic positions within the longitudinal sand dunes of the Gurbantunggut Desert, China. Journal of Arid Land, 9(3): 355–367.

Karchegani P M, Ayoubi S, Lu S G, et al. 2011. Use of magnetic measures to assess soil redistribution following deforestation in hilly region. Journal of Applied Geophysics, 75(2): 227–236.

Keshavarzi A, Tuffour H, Bagherzadeh A, et al. 2018. Spatial and fractal characterization of soil properties across soil depth in an agricultural field, Northeast Iran. Eurasian Journal of Soil Science, 7(2): 93–102.

Khaledian Y, Brevik E C, Pereira P, et al. 2017a. Modeling soil cation exchange capacity in multiple countries. Catena, 158: 194–200.

Khaledian Y, Kiani F, Ebrahimi S, et al. 2017b. Assessment and monitoring of soil degradation during land use change using multivariate analysis. Land Degradation & Development, 28(1): 128–141.

Khormali F, Ajami M, Ayoubi S, et al. 2009. Role of deforestation and hillslope position on soil quality attributes of loess-derived soils in Golestan province, Iran. Agriculture, Ecosystems & Environment, 134(3–4): 178–189.

Kravchenko A, Bullock D G. 1999. A comparative study of interpolation methods for mapping soil properties. Agronomy Journal, 91(3): 393–400.

Li J, Heap A D. 2011. A review of comparative studies of spatial interpolation methods in environmental sciences: Performance and impact factors. Ecological Informatics, 6(3–4): 228–241.

Li S S, Wang Q, Li L H. 2016. Interdecadal variations of pan-evaporation at the southern and northern slopes of the Tianshan Mountains, China. Journal of Arid Land, 8(6): 832–845.

Liao Y, Wu W L, Meng F Q, et al. 2015. Increase in soil organic carbon by agricultural intensification in northern China. Biogeosciences, 12(5): 1403–1413.

Malinowski E R. 2002. Factor Analysis in Chemistry. New York: John Wiley and Sons Press, 1–432.

Martínez-Hernández C, Rodrigo-Comino J, Romero-Díaz A. 2017. Impact of lithology and soil properties on abandoned dryland terraces during the early stages of soil erosion by water in Southeast Spain. Hydrological Processes, 31(17): 3095–3109.

Martínez-Murillo J F, Hueso-González P, Ruiz-Sinoga J D. 2017. Topsoil moisture mapping using geostatistical techniques under different Mediterranean climatic conditions. Science of The Total Environment, 595: 400–412.

Massart D L, Kaufman L. 1983. The Interpretation of Analytical Chemical Data by the Use of Cluster Analysis. New York: John Wiley and Sons Press, 237.

Mehnatkesh A, Ayoubi S, Jalalian A, et al. 2013. Relationships between soil depth and terrain attributes in a semi arid hilly region in western Iran. Journal of Mountain Science, 10(1): 163–172.

ASTER GDEM. 2009. ASTER Global Digital Elevation Model (ASTER GDEM). [2009-06-29]. http://www.jspacesystems.or.jp/ersdac/GDEM/E/.

Metternicht G I, Zinck J A, 2003. Remote sensing of soil salinity: potentials and constraints. Remote Sensing of Environment, 85: 1–20.

Minasny B, McBratney A B. 2006. A conditioned Latin hypercube method for sampling in the presence of ancillary information. Computers & geosciences, 32(9): 1378–1388.

Morin J, Van Winkel J. 1996. The effect of raindrop impact and sheet erosion on infiltration rate and crust formation. Soil Science Society of America Journal, 60(4): 1223–1227.

Nelson D W, Sommers E L. 1996. Total carbon, organic carbon, and organic matter. In: Sparks D L. Methods of Soil Analysis, Part 3, Chemical Methods. Madison: Soil Science Society of America Press, 961–1010.

Olaya V. 2004. A Gentle Introduction to SAGA GIS. Gottingen: The SAGA User Group Press, 1–216.

Orgill S E, Condon J R, Conyers M K, et al. 2017. Parent material and climate affect soil organic carbon fractions under pastures in south-eastern Australia. Soil Research, 55(8): 799–808.

Ortíz-Rodríguez A J, Borselli L, Sarocchi D. 2017. Flow connectivity in active volcanic areas: Use of index of connectivity in the assessment of lateral flow contribution to main streams. Catena, 157: 90–111.

Page A L, Miller R H, Keeney D R. 1982. Methods of Soil Analysis, Part 2. Chemical and Microbiological Properties. Madison: American Society of Agronomy, 1–1159.

Poeppl R E, Keesstra S D, Maroulis J. 2017. A conceptual connectivity framework for understanding geomorphic change in human-impacted fluvial systems. Geomorphology, 277: 237–250.

Ramos M C, Nacci S, Pla I. 2000. Soil sealing and its influence on erosion rates for some soils in the Mediterranean area. Soil Science, 165(5): 398–403.

Rodrigo-Comino J, Senciales J M, Cerdà A, et al. 2018. The multidisciplinary origin of soil geography: A review. Earth-Science Reviews, 177: 114–123.

Rosemary F, Indraratne S P, Weerasooriya R, et al. 2017. Exploring the spatial variability of soil properties in an Alfisol soil catena. Catena, 150: 53–61.

Ruggieri N, Castellano M, Capello M, et al. 2011. Seasonal and spatial variability of water quality parameters in the Port of Genoa, Italy, from 2000 to 2007. Marine Pollution Bulletin, 62(2): 340–349.

Sağir Ç, Kurtuluş B. 2017. Hydraulic head and groundwater $^{111}$Cd content interpolations using empirical Bayesian kriging (EBK) and geo-adaptive neuro-fuzzy inference system (geo-ANFIS). Water SA, 43(3): 509–519.

Samsonova V P, Blagoveshchenskii Y N, Meshalkina Y L. 2017. Use of empirical Bayesian kriging for revealing heterogeneities in the distribution of organic carbon on agricultural lands. Eurasian Soil Science, 50(3): 305–311.

Schjønning P, Munkholm L J, Elmholt S, et al. 2007. Organic matter and soil tilth in arable farming: Management makes a difference within 5–6 years. Agriculture, Ecosystems & Environment, 122(2): 157–172.

Shiri J, Keshavarzi A, Kisi O, et al. 2017. Modeling soil cation exchange capacity using soil parameters: Assessing the heuristic models. Computers and Electronics in Agriculture, 135: 242–251.

Shukla M K, Lal R, Ebinger M. 2006. Determining soil quality indicators by factor analysis. Soil and Tillage Research, 87(2): 194–204.

Singer M J, Shainberg I. 2004. Mineral soil surface crusts and wind and water erosion. Earth Surface Processes and Landforms, 29(9): 1065–1075.

Soil Survey Staff. 2014. Keys to Soil Taxonomy (12$^{th}$ ed.). Washington, D.C.: United States Department of Agriculture Natural Resources Conservation Service, 1–360.

Stavi I, Ungar E D, Lavee H, et al. 2008. Grazing-induced spatial variability of soil bulk density and content of moisture, organic carbon and calcium carbonate in a semi-arid rangeland. Catena, 75(3): 288–296.

Sulieman M, Saeed, I, Hassaballa A, et al. 2018. Modeling cation exchange capacity in multi geochronological-derived alluvium soils: An approach based on soil depth intervals. Catena, 167: 327–339.

Sun W Y, Zhu H H, Guo S L. 2015. Soil organic carbon as a function of land use and topography on the Loess Plateau of China. Ecological Engineering, 83: 249–257.

Taghizadeh-Mehrjardi R, Nabiollahi K, Kerry R. 2016. Digital mapping of soil organic carbon at multiple depths using different data mining techniques in Baneh region, Iran. Geoderma, 266: 98–110.

Tajik S, Ayoubi S, Nourbakhsh F, 2012. Prediction of soil enzymes activity by digital terrain analysis: comparing artificial neural network and multiple linear regression models. Environmental Engineering Science, 29(8): 798–806.

Tran C P, Bode R W, Smith A J, et al. 2010. Land-use proximity as a basis for assessing stream water quality in New York State (USA). Ecological Indicators, 10(3): 727–733.

U.S. Geological Survey. 2004. EarthExplorer Help Index: EarthExplorer Tutorial. https://earthexplorer.usgs.gov/.

Wang C, Zhao C Y, Xu Z L, et al. 2013. Effect of vegetation on soil water retention and storage in a semi-arid alpine forest catchment. Journal of Arid Land, 5(2): 207–219.

Wang J Q, Liu L C, Qiu X Q, et al. 2016. Contents of soil organic carbon and nitrogen in water-stable aggregates in abandoned agricultural lands in an arid ecosystem of Northwest China. Journal of Arid Land, 8(3): 350–363.

Wilford J, De Caritat P, Bui E. 2015. Modelling the abundance of soil calcium carbonate across Australia using geochemical survey data and environmental predictors. Geoderma, 259–260: 81–92.

Zeraatpishe M, Khormali F. 2012. Carbon stock and mineral factors controlling soil organic carbon in a climatic gradient, Golestan province. Journal of Soil Science and Plant Nutrition, 12(4): 637–654.

Zeraatpisheh M, Ayoubi S, Jafari A, et al. 2017. Comparing the efficiency of digital and conventional soil mapping to predict soil types in a semi-arid region in Iran. Geomorphology, 285: 186–204.

Zeraatpisheh M, Ayoubi S, Brungard C W, et al. 2019. Disaggregating and updating a legacy soil map using DSMART, fuzzy *c*-means and *k*-means clustering algorithms in Central Iran. Geoderma, 340: 249–258.